

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 10/041,054A  
Source: IFW/6  
Date Processed by STIC: 1/3/05

# ***ENTERED***



IFW16

## RAW SEQUENCE LISTING

DATE: 01/03/2005

PATENT APPLICATION: US/10/041,054A

TIME: 16:53:27

Input Set : A:\ORT1560NP.Subst.Seq.List.txt

Output Set: N:\CRF4\01032005\J041054A.raw

```

3 <110> APPLICANT: Darrow, Andrew
4      Qi, Jenson
5      Andrade-Gordon, Patricia
7 <120> TITLE OF INVENTION: DNA ENCODING THE HUMAN SERINE PROTEASE T
9 <130> FILE REFERENCE: ORT-1560
11 <140> CURRENT APPLICATION NUMBER: 10/041,054A
12 <141> CURRENT FILING DATE: 2002-01-07
14 <150> PRIOR APPLICATION NUMBER: 09/386,653
15 <151> PRIOR FILING DATE: 1999-08-31
17 <160> NUMBER OF SEQ ID NOS: 11
19 <170> SOFTWARE: PatentIn version 3.3
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 1110
23 <212> TYPE: DNA
24 <213> ORGANISM: Homo sapiens
26 <400> SEQUENCE: 1
27 gaccacggcc ctgcgccccca gccaggcctg aggacatgag gcggccggcg gcggtgccgc      60
29 tcctgctgct gctgtgtttt ggggtctcaga gggccaaggc agcaacagcc tgtggtcgcc      120
31 ccaggatgct gaaccgaatg gtgggcgggc aggacacgca ggaggcgag tggccctggc      180
33 aagtcagcat ccagcgcaac ggaagccact tctgcggggg cagcctcatc gcggagcagt      240
35 gggctcctgac ggctgcgcac tgcttcgcga acacctctga gacgtccctg taccaggtcc      300
37 tgctgggggc aaggcagcta gtgcagccgg gaccacacgc tatgtatgcc cgggtgaggc      360
39 aggtggagag caaccccctg taccagggca cggcctccag cgctgacgtg gccctggtgg      420
41 agctggaggg accagtgcct ttcaccaatt acatcctccc cgtgtgcctg cctgaccctt      480
43 cggatgatctt tgagacgggc atgaactgct gggtcactgg ctggggcagc cccagtgagg      540
45 aagacctcct gcccgaaccg cggatcctgc agaaactcgc tgtgcccctc atcgacacac      600
47 ccaagtgcaa cctgctctac agcaaagaca ccgagtttgg ctaccaaccc aaaaccatca      660
49 agaatgacat gctgtgcgcc ggcttcgagg agggcaagaa ggatgcctgc aagggcgact      720
51 cgggcggccc cctggtgtgc ctctggtggtc agtcgtggct gcaggcgggg gtgatcagct      780
53 ggggtgaggg ctgtgcccgc cagaaccgcc caggtgtcta catccgtgtc accgcccacc      840
55 acaactggat ccatcggatc atccccaac tgcagttcca gccagcgagg ttgggcggcc      900
57 agaagtgaga ccccgggggc caggagcccc ttgagcagag ctctgcaccc agcctgcccc      960
59 cccacaccat cctgctgggc ctcccagcgc tgctgttgca cctgtgagcc ccaccagact     1020
61 catttgtaaa tagcgtcctt tcctcccctc tcaaataccc ttattttatt tatgtttctc     1080
63 ccaataaaaa cccagcctgt gtgccagctg      1110
66 <210> SEQ ID NO: 2
67 <211> LENGTH: 20
68 <212> TYPE: DNA
69 <213> ORGANISM: Artificial
71 <220> FEATURE:
72 <223> OTHER INFORMATION: ProtT PCRTP-U PCR primer
74 <400> SEQUENCE: 2
75 gccaggcctg aggacatgag

```

20

## RAW SEQUENCE LISTING

DATE: 01/03/2005

PATENT APPLICATION: US/10/041,054A

TIME: 16:53:27

Input Set : A:\ORT1560NP.Subst.Seq.List.txt

Output Set: N:\CRF4\01032005\J041054A.raw

```

78 <210> SEQ ID NO: 3
79 <211> LENGTH: 20
80 <212> TYPE: DNA
81 <213> ORGANISM: Artificial
83 <220> FEATURE:
84 <223> OTHER INFORMATION: ProtT PC RTP-L PCR primer
86 <400> SEQUENCE: 3
87 tgcgctggat gctgacttgc                                20
90 <210> SEQ ID NO: 4
91 <211> LENGTH: 40
92 <212> TYPE: DNA
93 <213> ORGANISM: Artificial
95 <220> FEATURE:
96 <223> OTHER INFORMATION: ProtT PCTTP-PP primer
98 <400> SEQUENCE: 4
99 ccaggatgct gaaccgaatg gtgggcgggc aggacacgca          40
102 <210> SEQ ID NO: 5
103 <211> LENGTH: 30
104 <212> TYPE: DNA
105 <213> ORGANISM: Artificial
107 <220> FEATURE:
108 <223> OTHER INFORMATION: ProtT Xba-U PCR primer
110 <400> SEQUENCE: 5
111 aggatctaga ggagggcgag tggccctggc                    30
114 <210> SEQ ID NO: 6
115 <211> LENGTH: 30
116 <212> TYPE: DNA
117 <213> ORGANISM: Artificial
119 <220> FEATURE:
120 <223> OTHER INFORMATION: ProtT Xba-L PCR primer
122 <400> SEQUENCE: 6
123 ggggtctaga cttctggccg cccaacctcg                    30
126 <210> SEQ ID NO: 7
127 <211> LENGTH: 290
128 <212> TYPE: PRT
129 <213> ORGANISM: Homo sapiens
131 <400> SEQUENCE: 7
133 Met Arg Arg Pro Ala Ala Val Pro Leu Leu Leu Leu Cys Phe Gly
134 1          5          10          15
137 Ser Gln Arg Ala Lys Ala Ala Thr Ala Cys Gly Arg Pro Arg Met Leu
138          20          25          30
141 Asn Arg Met Val Gly Gly Gln Asp Thr Gln Glu Gly Glu Trp Pro Trp
142          35          40          45
145 Gln Val Ser Ile Gln Arg Asn Gly Ser His Phe Cys Gly Gly Ser Leu
146          50          55          60
149 Ile Ala Glu Gln Trp Val Leu Thr Ala Ala His Cys Phe Arg Asn Thr
150 65          70          75          80
153 Ser Glu Thr Ser Leu Tyr Gln Val Leu Leu Gly Ala Arg Gln Leu Val
154          85          90          95

```

## RAW SEQUENCE LISTING

DATE: 01/03/2005

PATENT APPLICATION: US/10/041,054A

TIME: 16:53:27

Input Set : A:\ORT1560NP.Subst.Seq.List.txt

Output Set: N:\CRF4\01032005\J041054A.raw

```

157 Gln Pro Gly Pro His Ala Met Tyr Ala Arg Val Arg Gln Val Glu Ser
158           100           105           110
161 Asn Pro Leu Tyr Gln Gly Thr Ala Ser Ser Ala Asp Val Ala Leu Val
162           115           120           125
165 Glu Leu Glu Ala Pro Val Pro Phe Thr Asn Tyr Ile Leu Pro Val Cys
166           130           135           140
169 Leu Pro Asp Pro Ser Val Ile Phe Glu Thr Gly Met Asn Cys Trp Val
170 145           150           155           160
173 Thr Gly Trp Gly Ser Pro Ser Glu Glu Asp Leu Leu Pro Glu Pro Arg
174           165           170           175
177 Ile Leu Gln Lys Leu Ala Val Pro Ile Ile Asp Thr Pro Lys Cys Asn
178           180           185           190
181 Leu Leu Tyr Ser Lys Asp Thr Glu Phe Gly Tyr Gln Pro Lys Thr Ile
182           195           200           205
185 Lys Asn Asp Met Leu Cys Ala Gly Phe Glu Glu Gly Lys Lys Asp Ala
186           210           215           220
189 Cys Lys Gly Asp Ser Gly Gly Pro Leu Val Cys Leu Val Gly Gln Ser
190 225           230           235           240
193 Trp Leu Gln Ala Gly Val Ile Ser Trp Gly Glu Gly Cys Ala Arg Gln
194           245           250           255
197 Asn Arg Pro Gly Val Tyr Ile Arg Val Thr Ala His His Asn Trp Ile
198           260           265           270
201 His Arg Ile Ile Pro Lys Leu Gln Phe Gln Pro Ala Arg Leu Gly Gly
202           275           280           285
205 Gln Lys
206           290

```

209 &lt;210&gt; SEQ ID NO: 8

210 &lt;211&gt; LENGTH: 1130

211 &lt;212&gt; TYPE: DNA

212 &lt;213&gt; ORGANISM: Artificial

214 &lt;220&gt; FEATURE:

215 &lt;223&gt; OTHER INFORMATION: PFEK-PROTT-HIS fusion protein nucleic acid sequence

217 &lt;400&gt; SEQUENCE: 8

```

218 gaattcacca ccattggacag caaaggttcg tgcagaaat cccgcctgct cctgctgctg      60
220 gtggtgtcaa atctactctt gtgccagggt gtggtctccg actacaagga cgacgacgac      120
222 gtggacgcgg ccgctcttgc tgcccccttt gatgatgatg acaagatcgt tgggggctat      180
224 gctctagagg agggcgagtg gccctggcaa gtcagcatcc agcgcaacgg aagccacttc      240
226 tgcgggggca gcctcatcgc ggagcagtggt gtcctgacgg ctgcgcactg cttccgcaac      300
228 acctctgaga cgctccctgta ccagggtcctg ctgggggcaa ggcagctagt gcagccggga      360
230 ccacacgcta tgtatgcccg ggtgaggcag gtggagagca acccctgta ccagggcacg      420
232 gcctccagcg ctgacgtggc cctggtggag ctggaggcac cagtgcctt caccaattac      480
234 atcctccccg tgtgcttgc tgacctctcg gtgatctttg agacgggcat gaactgctgg      540
236 gtcactggct ggggcagccc cagtgaggaa gacctctgc ccgaaccgcg gatcctgcag      600
238 aaactcgctg tgcccatcat cgacacacc aagtgaacc tgctctacag caaagacacc      660
240 gagtttggt accaaccaca aaccatcaag aatgacatgc tgtgcgccg cttcgaggag      720
242 ggcaagaagg atgcctgcaa gggcgactcg ggcggccccc tgggtgtgct cgtgggtcag      780
244 tcgtggctgc aggcgggggt gatcagctgg ggtgagggt gtgcccgcca gaaccgcca      840
246 ggtgtctaca tccgtgtcac cgcccaccac aactggatcc atcgatcat ccccaaactg      900
248 cagttccagc cagcgaggtt gggcgccag aagtctagac atcaccatca ccatcactag      960

```

## RAW SEQUENCE LISTING

DATE: 01/03/2005

PATENT APPLICATION: US/10/041,054A

TIME: 16:53:27

Input Set : A:\ORT1560NP.Subst.Seq.List.txt

Output Set: N:\CRF4\01032005\J041054A.raw

```

250 cggccgcttc ccttttagtga ggggtaatgc ttcgagcaga catgataaga tacattgatg 1020
252 agtttgacaa aaccacaact agaatgcagt gaaaaaaatg ctttatttgt gaaatttgtg 1080
254 atgctattgc tttatttgta accattataa gctgcaataa acaagttgac 1130
257 <210> SEQ ID NO: 9
258 <211> LENGTH: 315
259 <212> TYPE: PRT
260 <213> ORGANISM: Artificial
262 <220> FEATURE:
263 <223> OTHER INFORMATION: PFEK-PROTT-HIS fusion protein amino acid sequence
265 <400> SEQUENCE: 9
267 Met Asp Ser Lys Gly Ser Ser Gln Lys Ser Arg Leu Leu Leu Leu Leu
268 1 5 10 15
271 Val Val Ser Asn Leu Leu Leu Cys Gln Gly Val Val Ser Asp Tyr Lys
272 20 25 30
275 Asp Asp Asp Asp Val Asp Ala Ala Ala Leu Ala Ala Pro Phe Asp Asp
276 35 40 45
279 Asp Asp Lys Ile Val Gly Gly Tyr Ala Leu Glu Glu Gly Glu Trp Pro
280 50 55 60
283 Trp Gln Val Ser Ile Gln Arg Asn Gly Ser His Phe Cys Gly Gly Ser
284 65 70 75 80
287 Leu Ile Ala Glu Gln Trp Val Leu Thr Ala Ala His Cys Phe Arg Asn
288 85 90 95
291 Thr Ser Glu Thr Ser Leu Tyr Gln Val Leu Leu Gly Ala Arg Gln Leu
292 100 105 110
295 Val Gln Pro Gly Pro His Ala Met Tyr Ala Arg Val Arg Gln Val Glu
296 115 120 125
299 Ser Asn Pro Leu Tyr Gln Gly Thr Ala Ser Ser Ala Asp Val Ala Leu
300 130 135 140
303 Val Glu Leu Glu Ala Pro Val Pro Phe Thr Asn Tyr Ile Leu Pro Val
304 145 150 155 160
307 Cys Leu Pro Asp Pro Ser Val Ile Phe Glu Thr Gly Met Asn Cys Trp
308 165 170 175
311 Val Thr Gly Trp Gly Ser Pro Ser Glu Glu Asp Leu Leu Pro Glu Pro
312 180 185 190
315 Arg Ile Leu Gln Lys Leu Ala Val Pro Ile Ile Asp Thr Pro Lys Cys
316 195 200 205
319 Asn Leu Leu Tyr Ser Lys Asp Thr Glu Phe Gly Tyr Gln Pro Lys Thr
320 210 215 220
323 Ile Lys Asn Asp Met Leu Cys Ala Gly Phe Glu Glu Gly Lys Lys Asp
324 225 230 235 240
327 Ala Cys Lys Gly Asp Ser Gly Gly Pro Leu Val Cys Leu Val Gly Gln
328 245 250 255
331 Ser Trp Leu Gln Ala Gly Val Ile Ser Trp Gly Glu Gly Cys Ala Arg
332 260 265 270
335 Gln Asn Arg Pro Gly Val Tyr Ile Arg Val Thr Ala His His Asn Trp
336 275 280 285
339 Ile His Arg Ile Ile Pro Lys Leu Gln Phe Gln Pro Ala Arg Leu Gly
340 290 295 300
343 Gly Gln Lys Ser Arg His His His His His His

```

## RAW SEQUENCE LISTING

DATE: 01/03/2005

PATENT APPLICATION: US/10/041,054A

TIME: 16:53:27

Input Set : A:\ORT1560NP.Subst.Seq.List.txt

Output Set: N:\CRF4\01032005\J041054A.raw

```
344 305          310          315
347 <210> SEQ ID NO: 10
348 <211> LENGTH: 4
349 <212> TYPE: PRT
350 <213> ORGANISM: Artificial
352 <220> FEATURE:
353 <223> OTHER INFORMATION: Chromogenic substrate 5
356 <220> FEATURE:
357 <221> NAME/KEY: MISC_FEATURE
358 <222> LOCATION: (1)..(1)
359 <223> OTHER INFORMATION: N-Succinyl-alanine
361 <220> FEATURE:
362 <221> NAME/KEY: MISC_FEATURE
363 <222> LOCATION: (4)..(4)
364 <223> OTHER INFORMATION: Phe-p-nitroanilide
366 <400> SEQUENCE: 10
W--> 368 Xaa Ala Pro Xaa
369 1
372 <210> SEQ ID NO: 11
373 <211> LENGTH: 4
374 <212> TYPE: PRT
375 <213> ORGANISM: Artificial
377 <220> FEATURE:
378 <223> OTHER INFORMATION: Chromogenic substrate 6
381 <220> FEATURE:
382 <221> NAME/KEY: MISC_FEATURE
383 <222> LOCATION: (1)..(1)
384 <223> OTHER INFORMATION: N-(methoxysuccinyl)-Ala
386 <220> FEATURE:
387 <221> NAME/KEY: MISC_FEATURE
388 <222> LOCATION: (4)..(4)
389 <223> OTHER INFORMATION: Val-p-nitroanilide
391 <400> SEQUENCE: 11
W--> 393 Xaa Ala Pro Xaa
394 1
```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 01/03/2005  
PATENT APPLICATION: US/10/041,054A      TIME: 16:53:28

Input Set : A:\ORT1560NP.Subst.Seq.List.txt  
Output Set: N:\CRF4\01032005\J041054A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:10; Xaa Pos. 1,4  
Seq#:11; Xaa Pos. 1,4

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:2,3,4,5,6,8,9,10,11

## VERIFICATION SUMMARY

DATE: 01/03/2005

PATENT APPLICATION: US/10/041,054A

TIME: 16:53:28

Input Set : A:\ORT1560NP.Subst.Seq.List.txt

Output Set: N:\CRF4\01032005\J041054A.raw

L:368 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:0  
L:393 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0